

## The word '*grid*' is (over)used a lot (HYPE)

- Oracle databases ;(
- cluster computing
- cycle scavenging
- “If a customer calls it a ‘grid’, then it is a grid”
- **cross-domain resource and data sharing**

## Is there a clear definition?

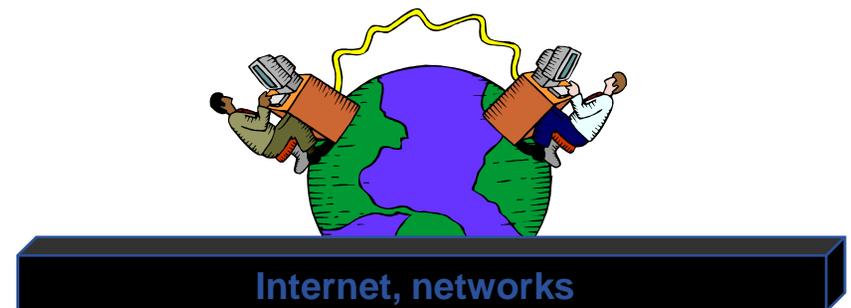
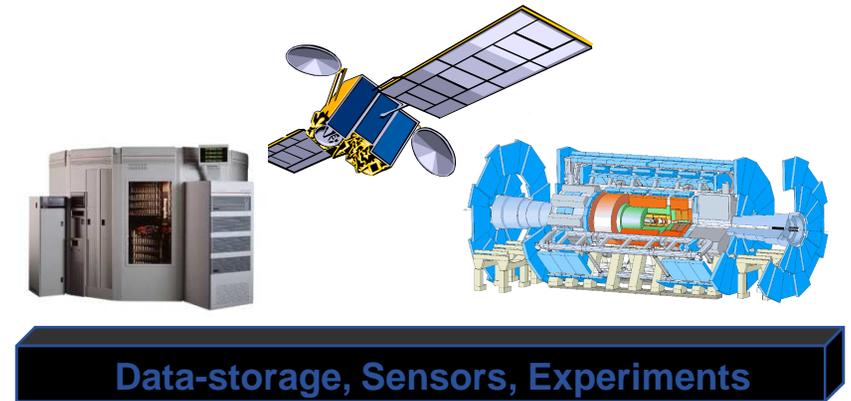
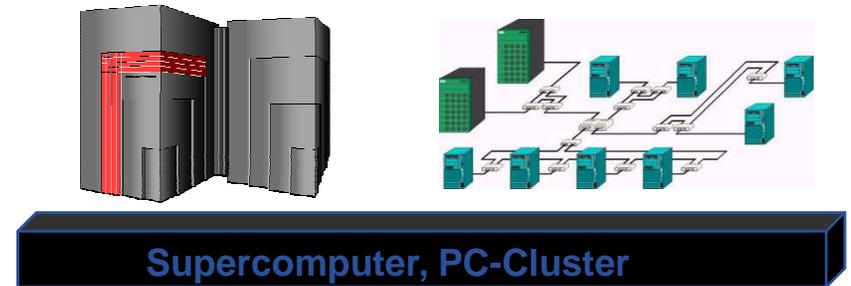
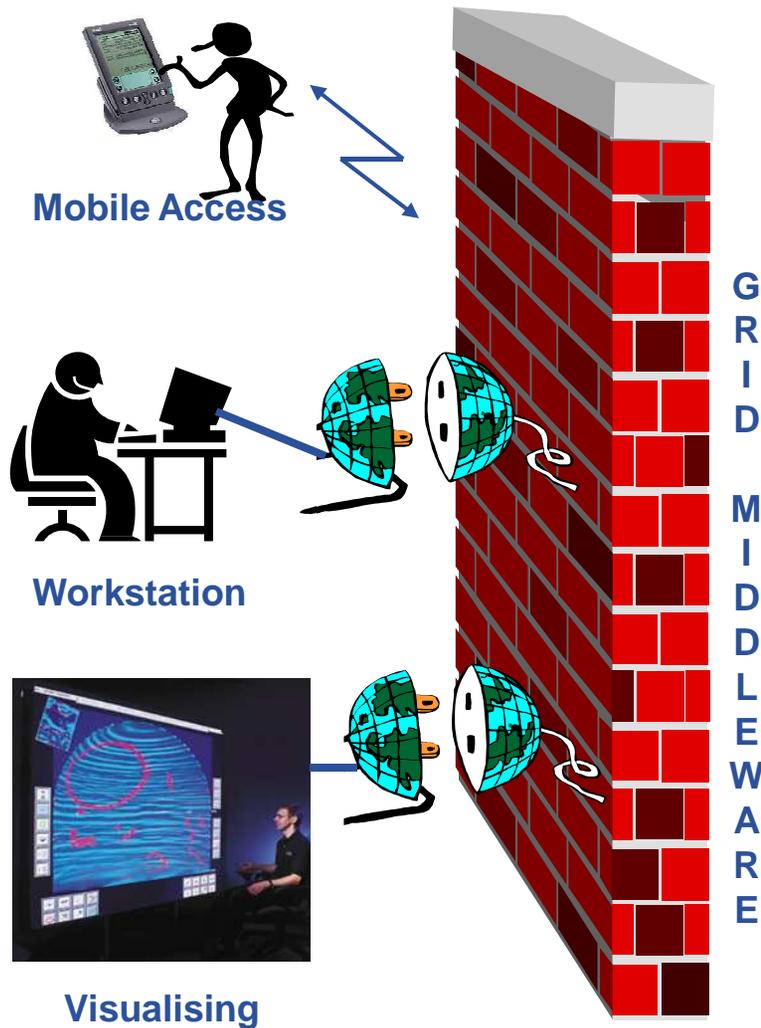
- Coördinate resources not under a central controle
- The use of standards, open and generic protocols & interfaces
- Delivering a non-trivial amount of collective services

### When do you need a grid?

- More then one computer
- More then one use (sharing)
- More then one location (collaborating)
- More then one company/devison
- More then one community

In general: More then ONE.

# The Grid metaphor





Enabling Grids for E-scienceE

# Introduction GRID Tutorial

Maurice Bouwhuis

SARA

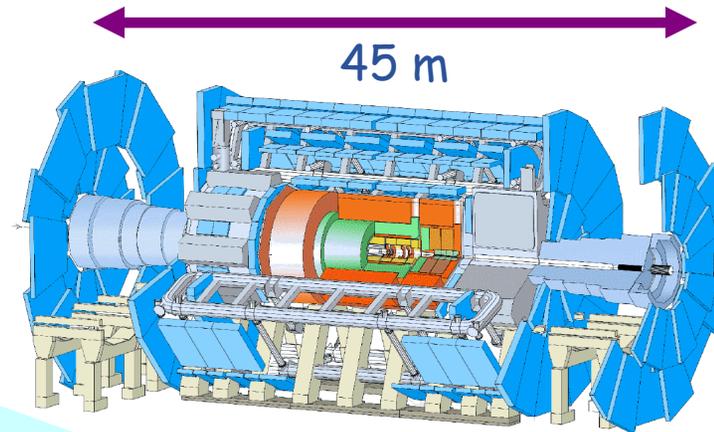
[www.eu-egee.org](http://www.eu-egee.org)



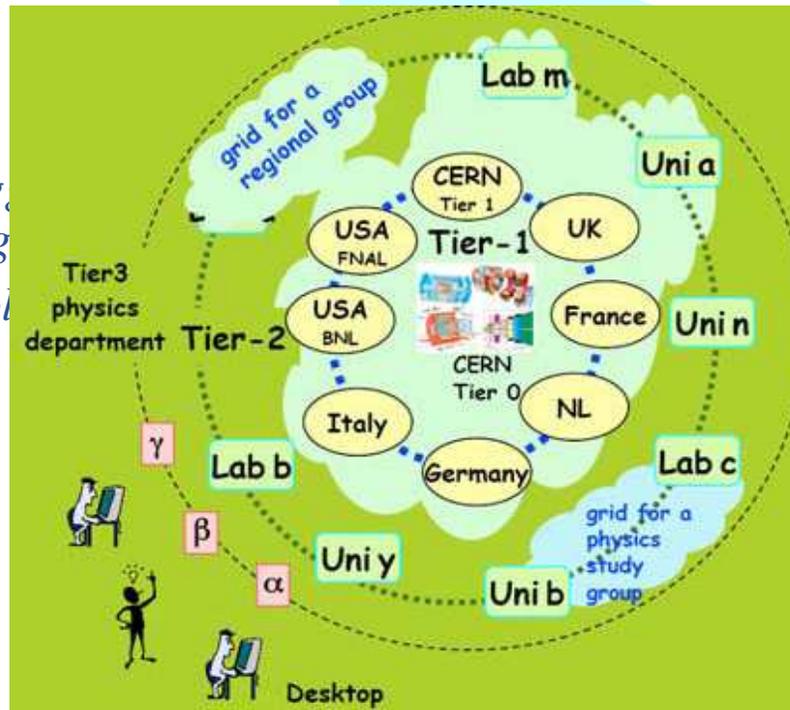


## ATLAS

*One of the four  
LHC detectors*

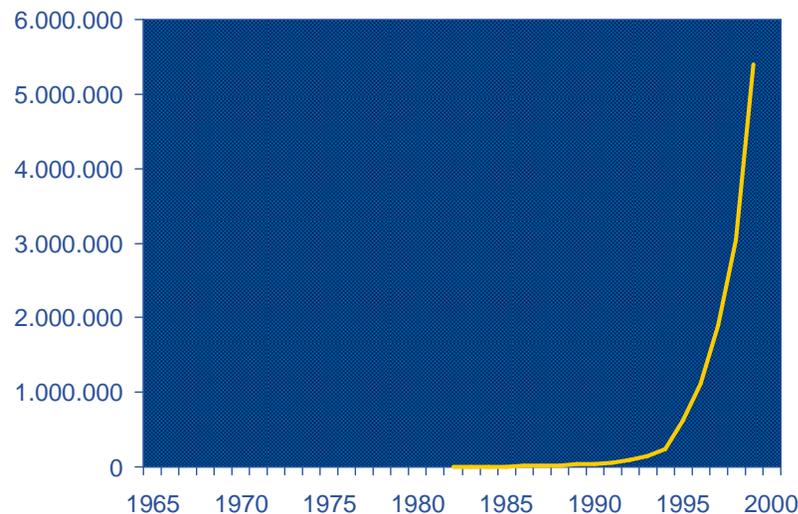


*online system  
multi-level trigger  
filter out background  
reduce data volume*



In NL:  
 -~3000 CPU  
 -3.000.000 GB disk  
 -3.000.000 GB tape per jaar

- Autometisering en increased resolution → more data AND more complex data
- Many different sources
- Hypothesis driven → data driven



Over 5 million  
sequence entries in  
GenBank

Over 3 billion  
bases from  
41,000 species

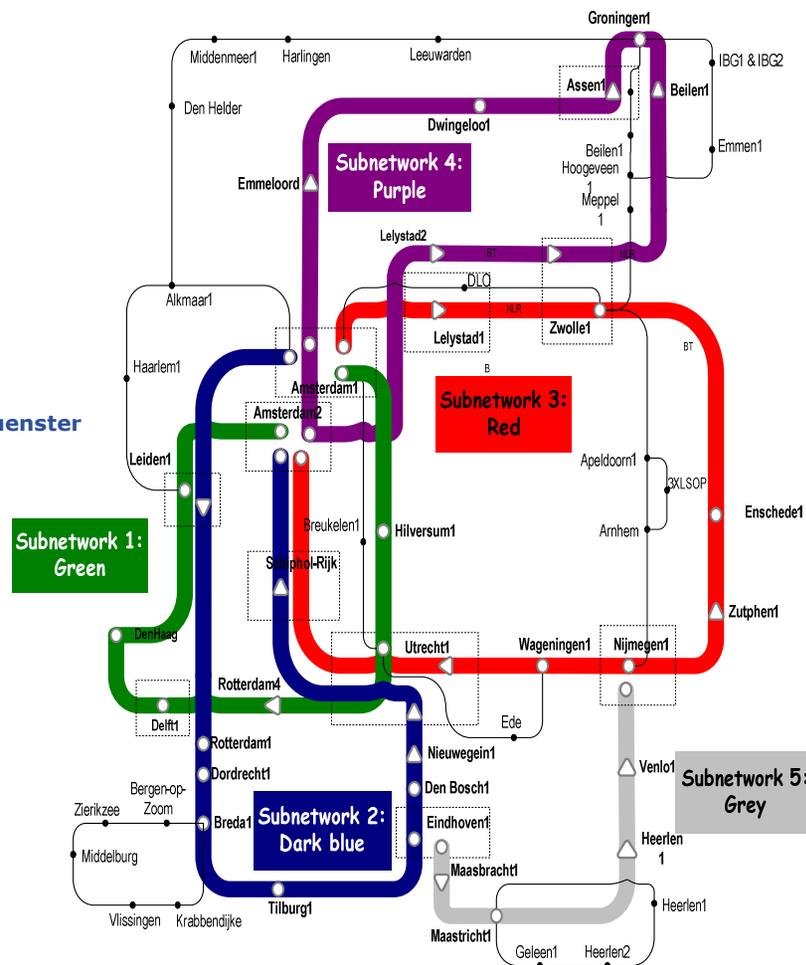
- **Astronomy (LOFAR et al)**
- **Climate research**
- **Earth observation**
- **Alpha en Gamma sciences**
  - Storage and Long Term archiving
  - Analysis of digital files
- **Ecology**
- **Food and health**
- **Medical instrumentation design**
- **Archaeology**
- .....

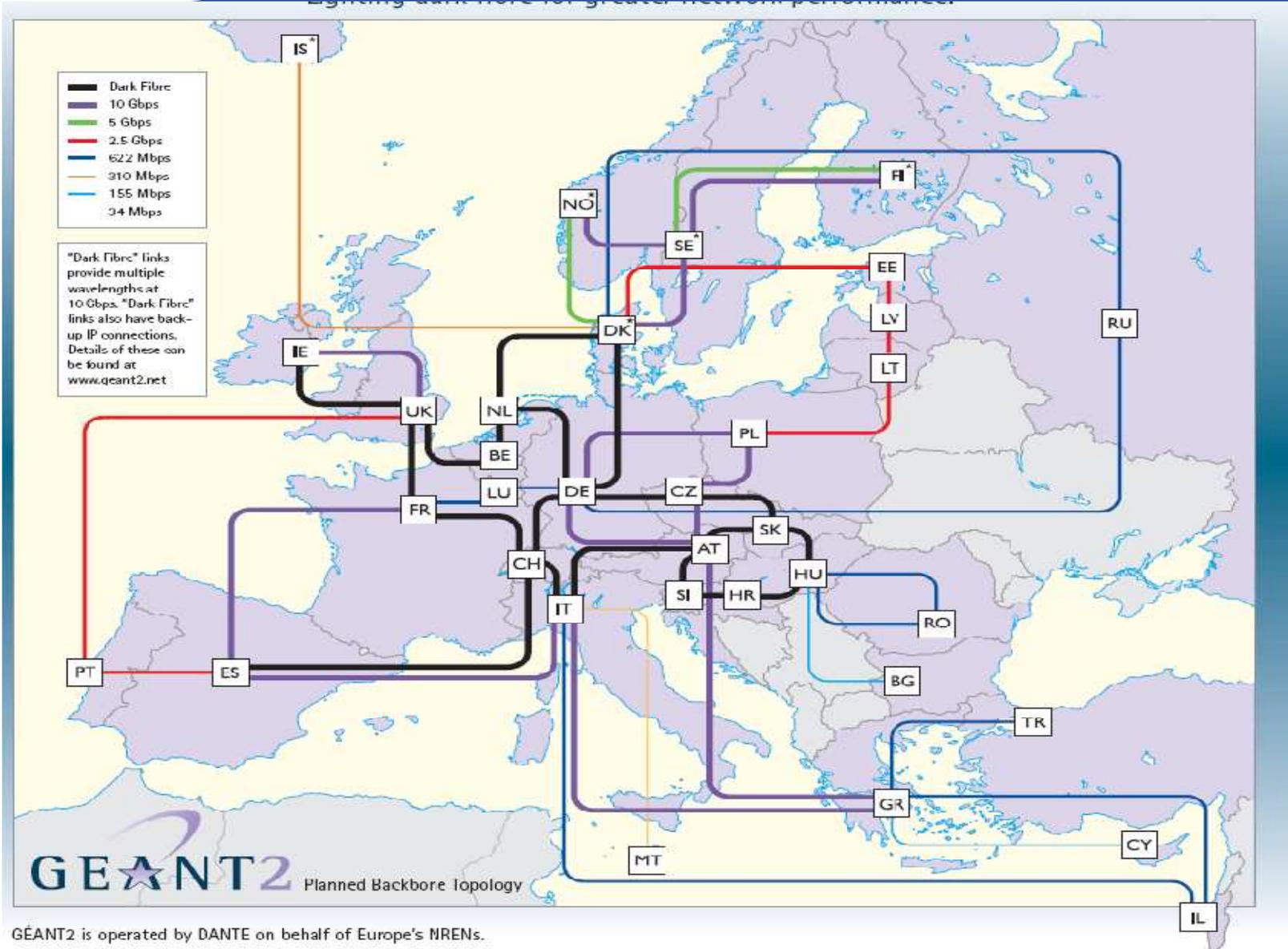
Where are you from

- **So what is happening today?**
  - Scale! Grid infrastructures operate worldwide
    - International infrastructures - EGEE, DEISA, Nordugrid, OSG, TeraGrid
    - National – NAREGI (Japan), UK-eScience, D-Grid, NLGrid
  - Interoperability – availability of middleware – Globus toolkit, UNICORE, NAREGI, schedulers

- **Some basic requirements for a grid infrastructure**
  - Transparent user administration – single sign on (single grid identity), authorisation and accounting based on grid identity – AAA facilities
  - Job scheduling – which can handle different environments
  - Global data access
  - Global information services – job information, data information, resource information
- **Interoperability!**
  - Standards needed for federation of infrastructures – GGF, IETF.....

- **Developments in network connectivity (high bandwidths) and tools play an important role**
  - 10 Gbps WAN links available today, both shared links and dedicated lightpaths (based on lambda technology)
  - 1 Gbps network adapters are commodity items on systems today and 10GE adapters available

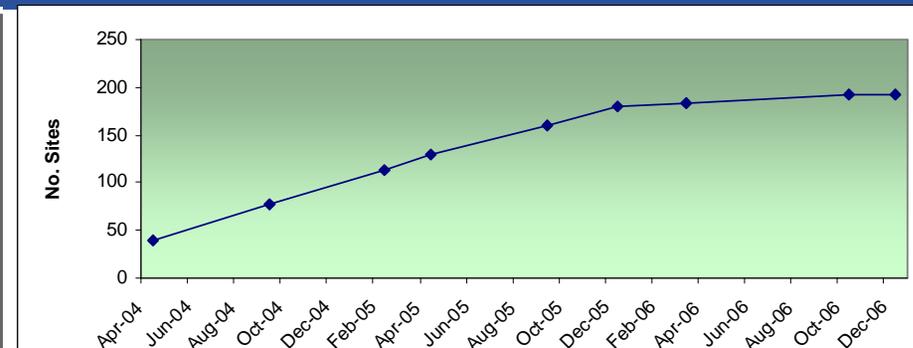
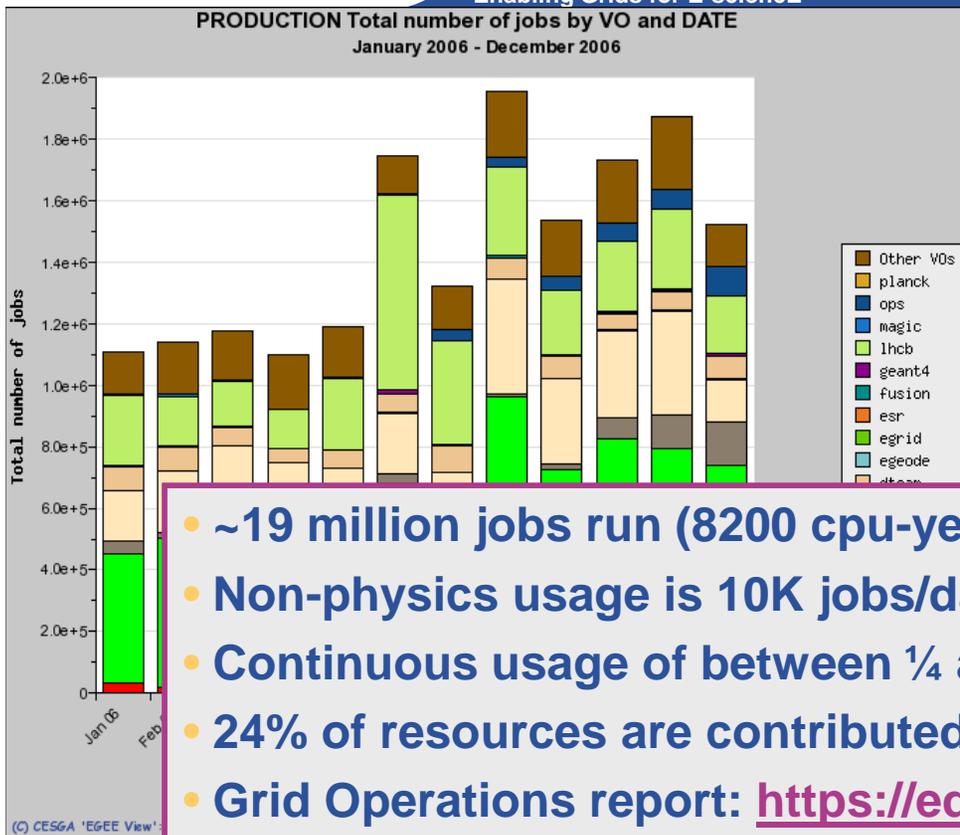




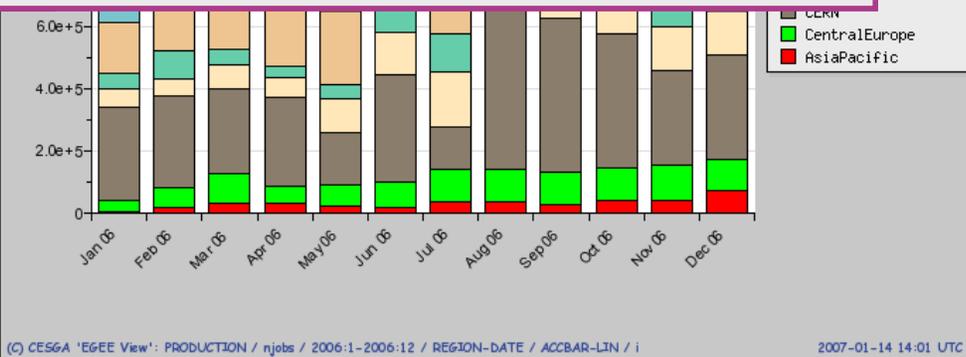
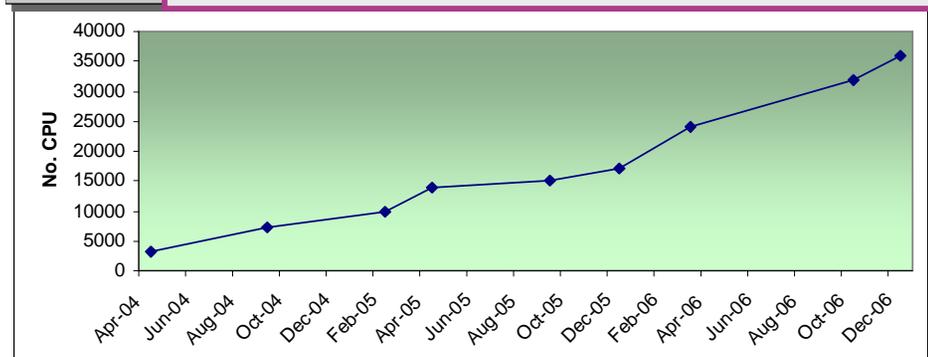


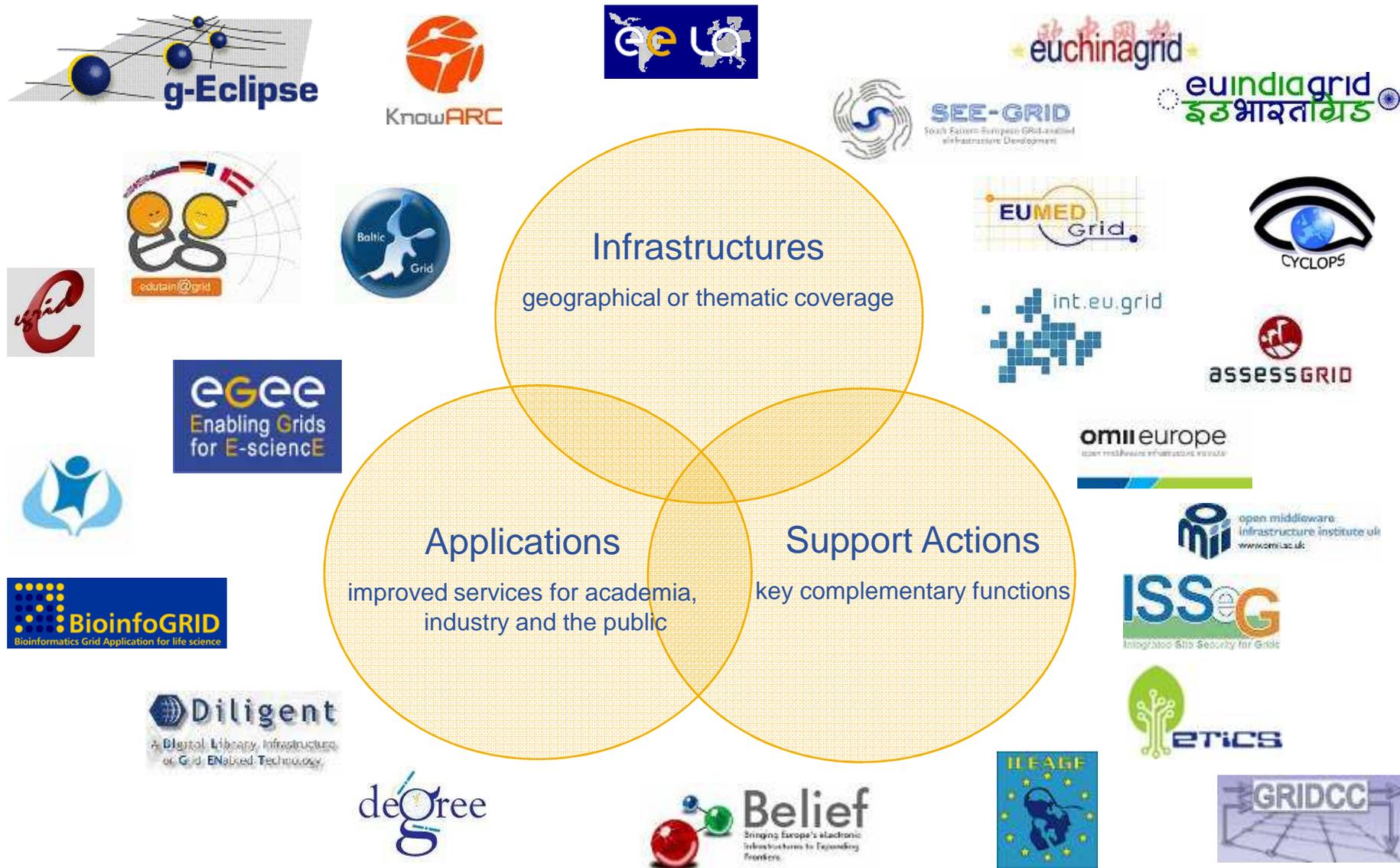
- **EGEE**
  - 1 April 2004 – 31 March 2006
  - 71 partners in 27 countries, federated in regional Grids
  - Operation of a pan European production Grid
- **EGEE-II**
  - 1 April 2006 – 31 March 2008
  - Expanded consortium
    - 91 partners
    - 11 Joint Research Units
  - Natural continuation of EGEE
  - Emphasis on providing production-level infrastructure
    - increased support for applications
    - interoperation with other Grid infrastructures
    - more involvement from Industry

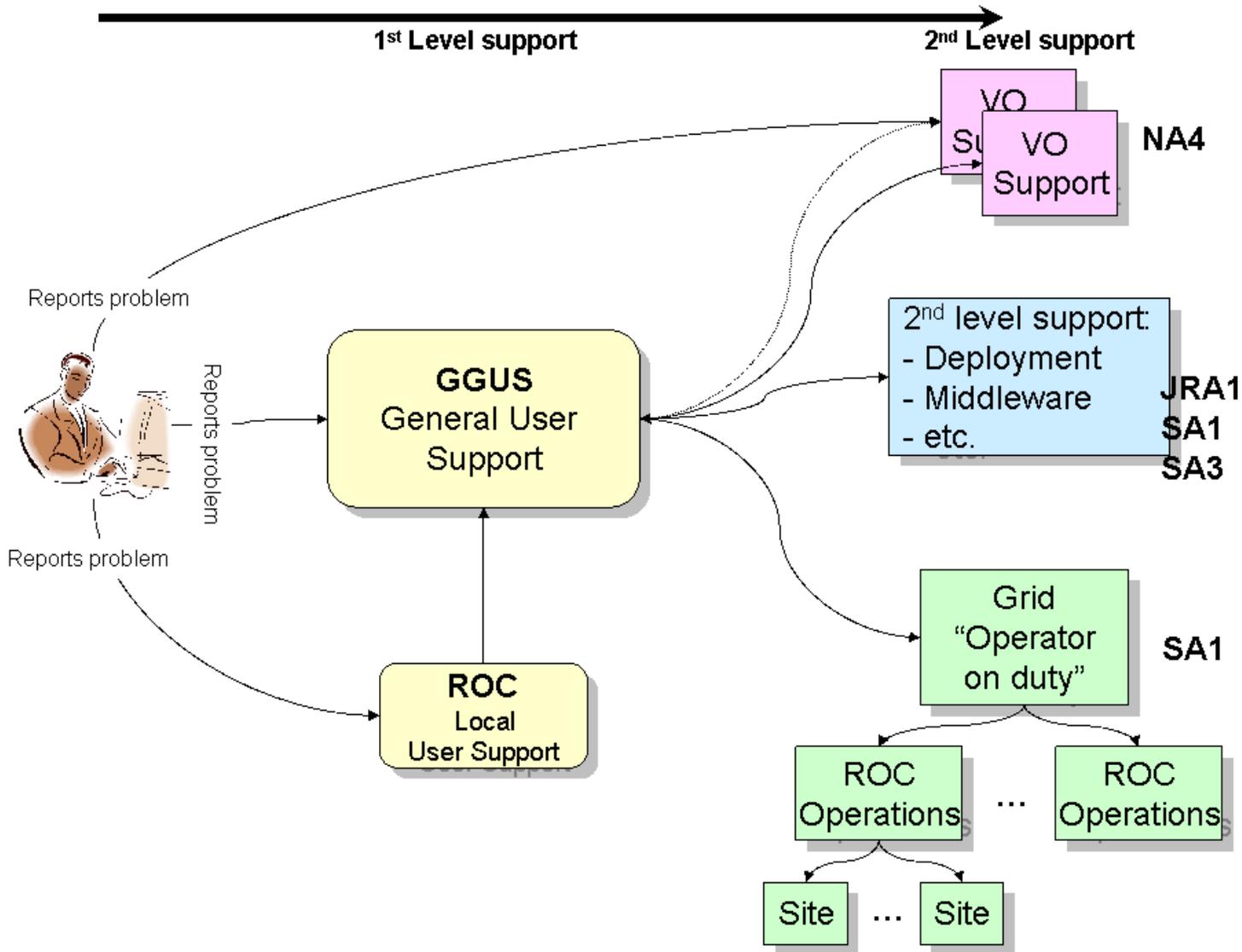
- **Mission**
  - Manage and operate production e-Infrastructure open to all user communities and service providers
  - Contribute to Grid standardisation and policy efforts
- **Infrastructure operation**
  - Currently include ~200 sites across 39 countries
  - Continuous monitoring of Grid services in a distributed global infrastructure
  - Automated site configuration/management
- **Future**
  - Expand on interoperability with related infrastructures



- ~19 million jobs run (8200 cpu-years, ~50K jobs/day) in 2006
- Non-physics usage is 10K jobs/day (same as whole of EGEE in 2005)
- Continuous usage of between 1/4 and 1/3 of the available resources
- 24% of resources are contributed by groups external to the project
- Grid Operations report: <https://edms.cern.ch/document/726140>

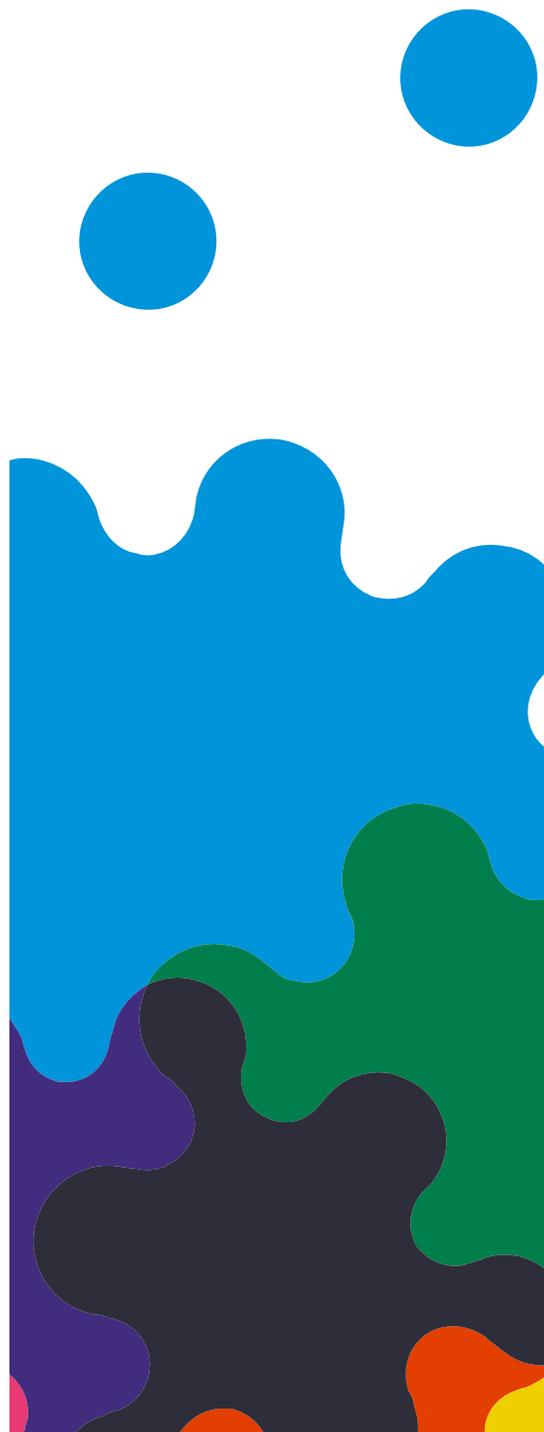






- **NE website:** <http://www.egee-ne.org/operations>
- **User support: contact user support at local site or mail to support@egee-ne.org**
  - NE uses a ticketing system monitored by different partners from our region. In NL NIKHEF, RC-RuG, SARA responsible.
  - Tickets from GGUS are also imported in the NE system
- **Application support – NA4 activity. In NL RC-RuG, SARA**





# **BiG** Grid

the dutch e-science grid



# BiG Grid

- **Strengthen existing National Grid infrastructure in Netherlands (NL-GRID by NCF)**
- **Sudsidy of 28 M€ for hardware and peopleware (expertise and support)**
- **Core partners**
  - NCF
  - Nikhef (High Energy Physics)
  - NBIC (BioInformaitcs)
- **Central and Distributed facilities**

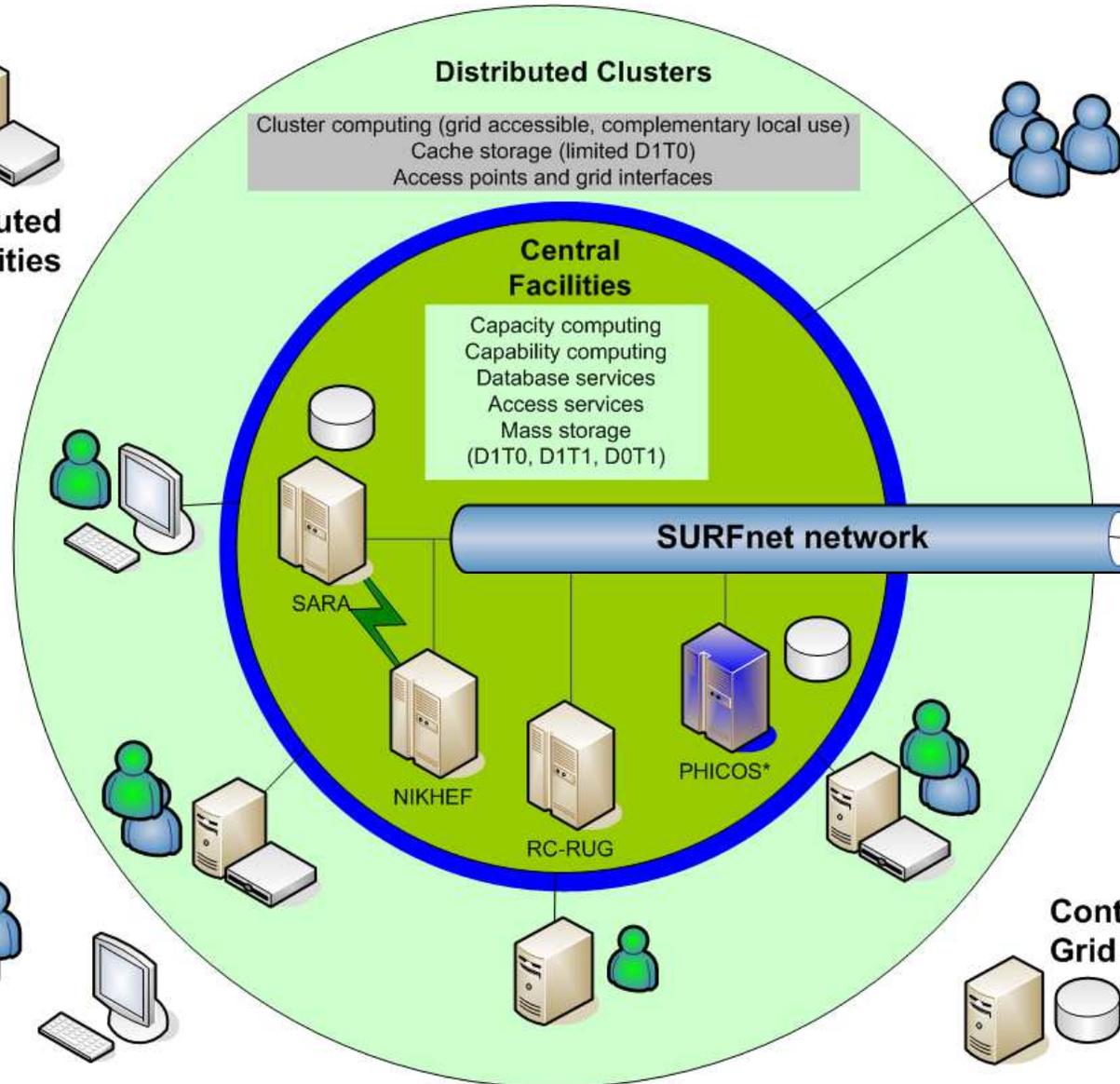


**BiG Grid**

the dutch e-science grid



**Contributed  
Grid Facilities**



**Contributed  
Grid Facilities**



**BiG Grid**  
the dutch e-science grid



# Infrastrucure



Central Facilities (operational)



Life Science Grid (operational)



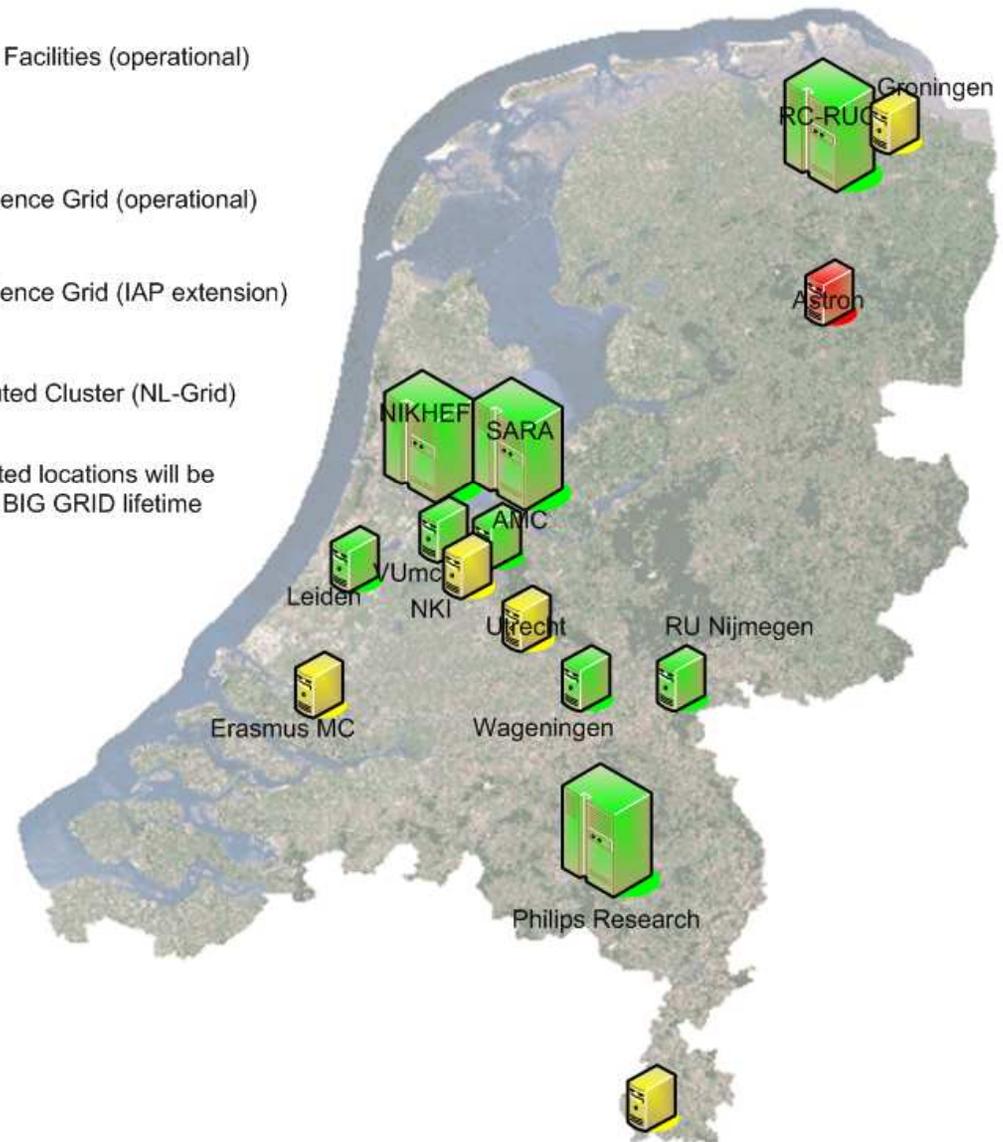
Life Science Grid (IAP extension)



Distributed Cluster (NL-Grid)

Additional distributed locations will be specified over the BIG GRID lifetime

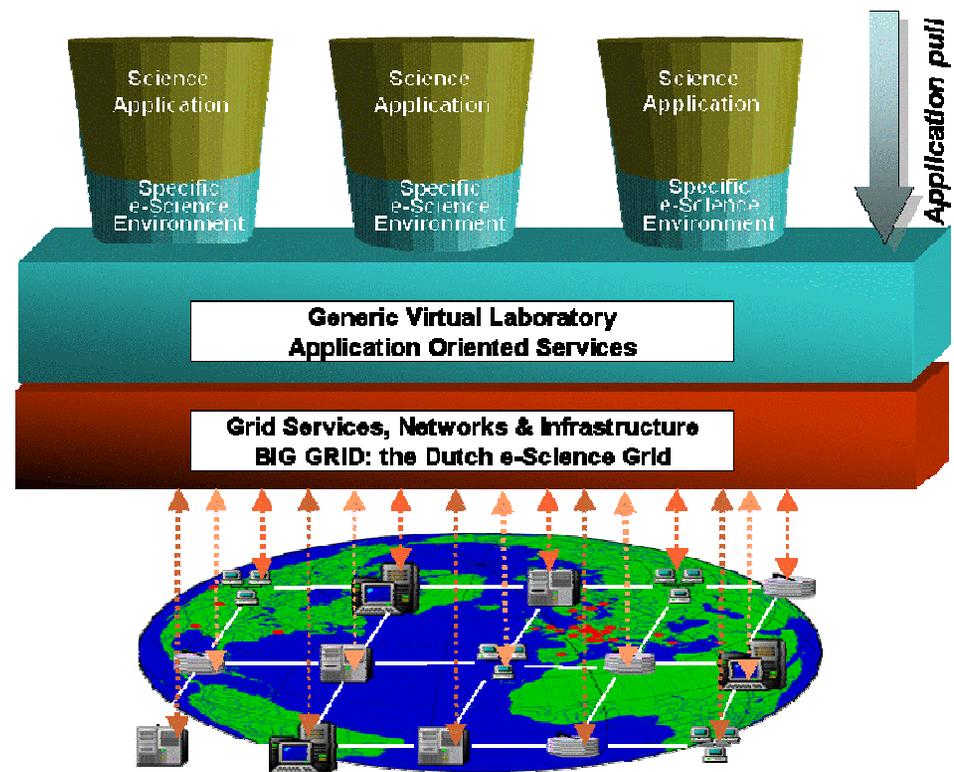
**O(5000) CPU**  
**O(10) PB disk storage**  
**O(20) PB tape storage**  
**O(10) Life Science Grid clusters**



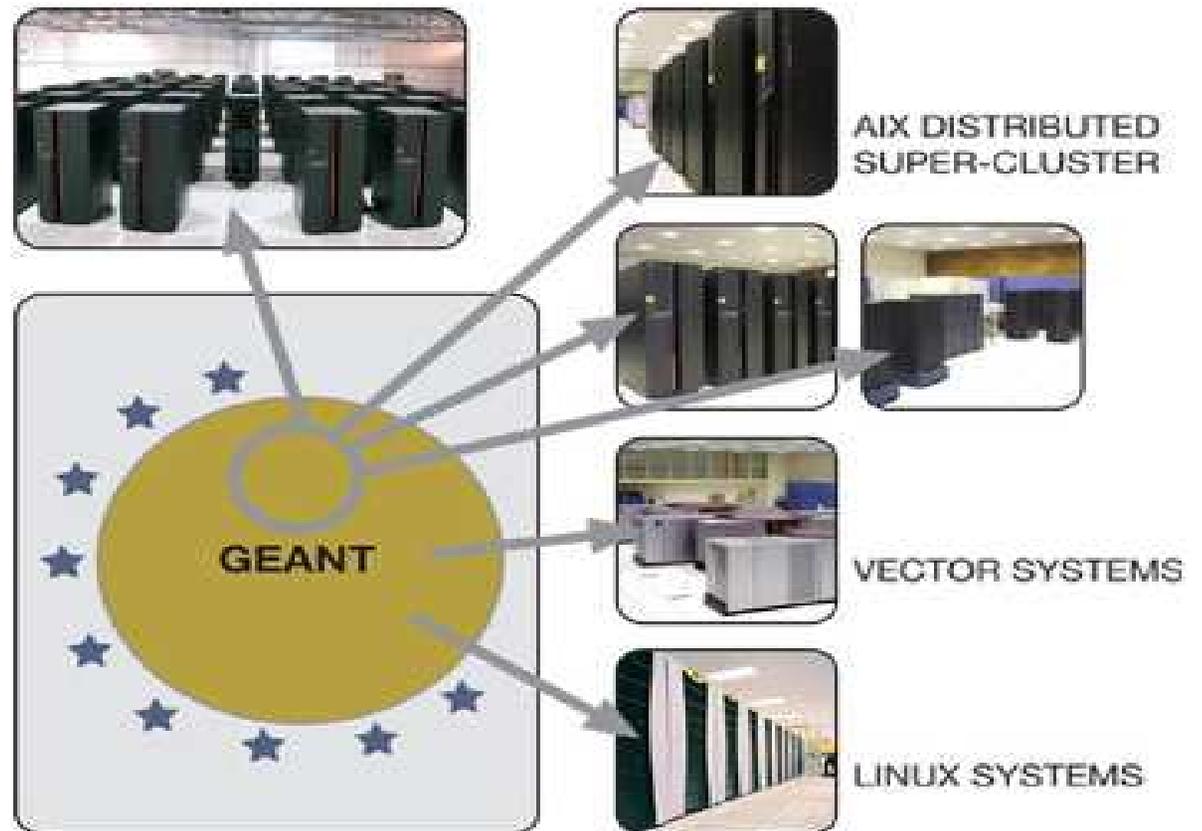
**BiG Grid**  
the dutch e-science

# Combination of 'push' and 'pull'

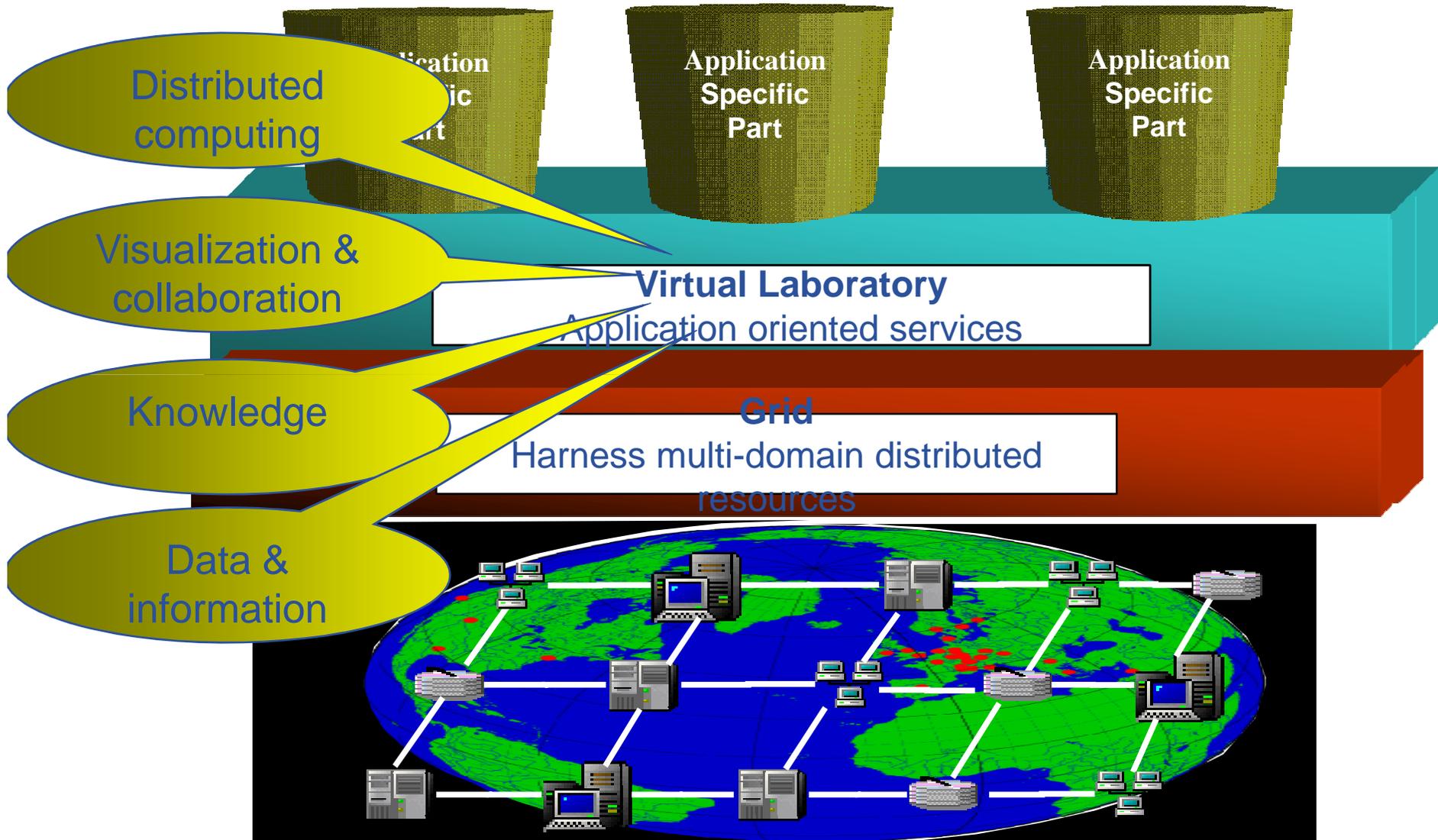
- **Application support:**
  - expertise: Application Domain Analysts
  - help desk and operations centre
- **Uniform software suite → Collaboration**
  - Standarts:
    - Open Grid Forum
  - EGEE: 'production' grid (40 disciplines)
  - BSIK VL-E project



- European super-computing grid
- Shared global file system
- Job migration
- Co-scheduling



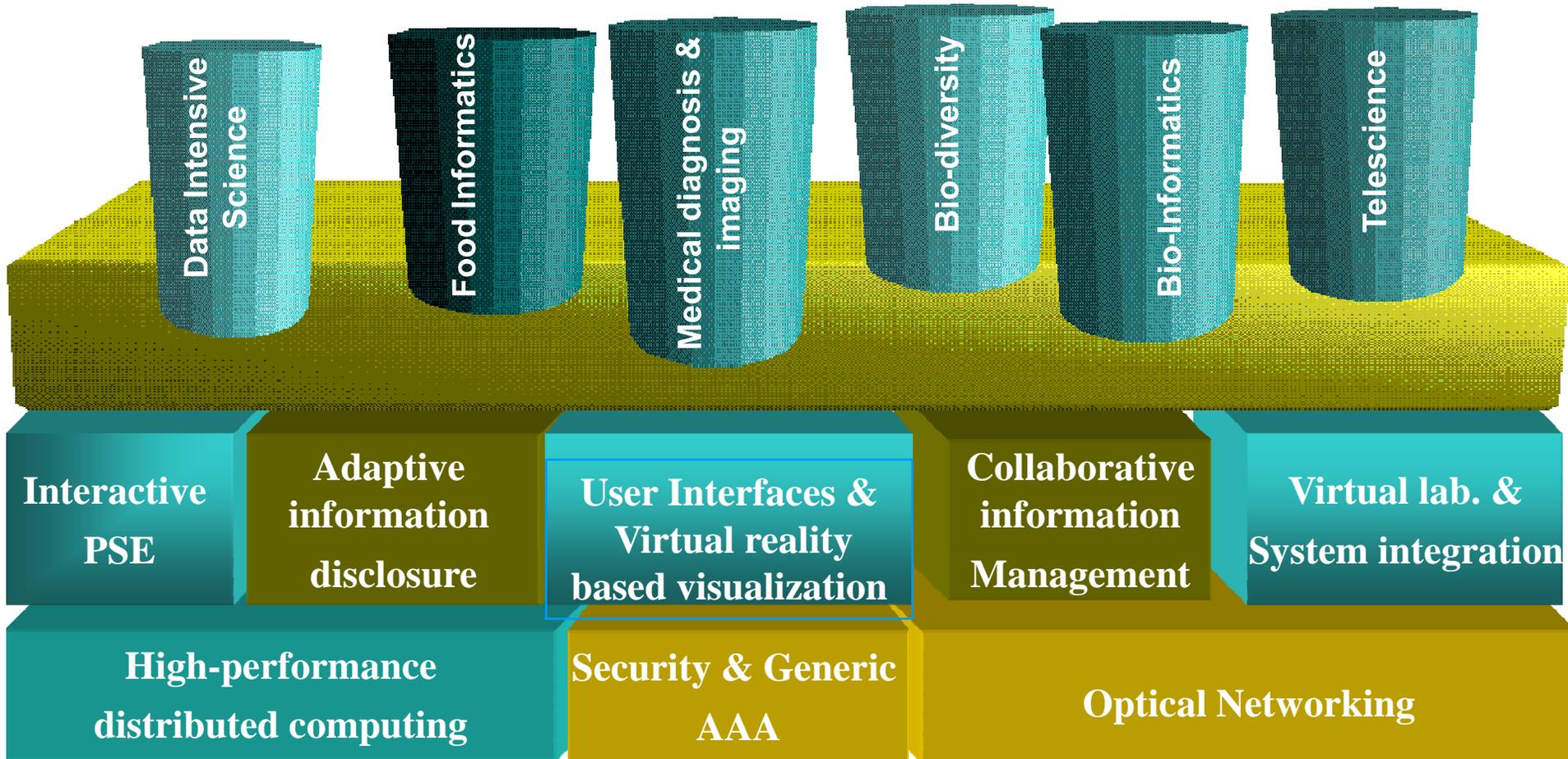
# Virtual Laboratories



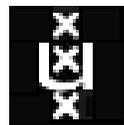


Enabling Grids for E-science

# Virtual Laboratory for e-Science



- 40 M€ (20 M€ BSIK funding)
- 2004 - 2008
- 20 partners
- Academic - Industrial



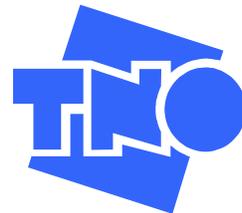
UNIVERSITEIT VAN AMSTERDAM



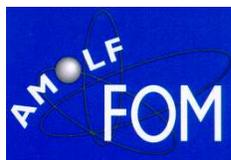
Unilever



vrije Universiteit



WAGENINGEN  
CENTRE  
FOR FOOD  
SCIENCES



- **Day 1**
  - Authentication
  - Job Submission
  - Tools for Collaboration by SURFnet
  - Drinks
  
- **Day 2**
  - Data handling
  - User Scenarios (VI-e and DANS)

Tutorial is free thanks to support of our sponsors



• **Gridforum.nl**



• **Netherlands Center for Bioinformatics**



• **NCF**



• **SURFnet**